Allophonic Cues to Syllabification

Andries W. Coetzee and Kevin McGowan

CUNY Syllable Conference

January 2008

Hidden structure

- ❖ Speech is just a continuous, sequential stream of noise.
  - The acoustic signal does not contain syllables, feet, moras, etc.
  - But so much of phonology is based on these concepts.
  - So, if they are real, then the mind must impose these structures on the acoustic signal in the process of mapping noise to linguistic percept.

- ❖ How can we determine whether these things are real?
  - Since they are not physically present in the acoustic signal, we cannot simply do an acoustic analysis.
  - But these hidden structures are part of the processing that listeners perform on the acoustic signal.
  - So, maybe we can look for evidence that the linguistic percept that listeners form show signs of these hidden structures being imposed on the acoustic signal – i.e. a behavioral data.

Allophonic distributions

❖ Some allophonic distribution rules make reference to syllable structure.

- /l/-allophones in English

$$/l/ \rightarrow \left\{ \begin{array}{ll} [\text{ɫ}] & \text{in coda} \\ [\text{l}] & \text{in onset} \end{array} \right\}$$

milk	[mɪɫk]
late	[let]

- Voiceless stops in English (simplified)

$$/\text{p, t, k}/ \rightarrow \left\{ \begin{array}{ll} [\text{p}^h, \text{t}^h, \text{k}^h] & / \ _{\sigma}[\ \_\_ \\ [\text{p, t, k}] & / \ _{\sigma}[\text{s}\ \_\_ \end{array} \right\}$$

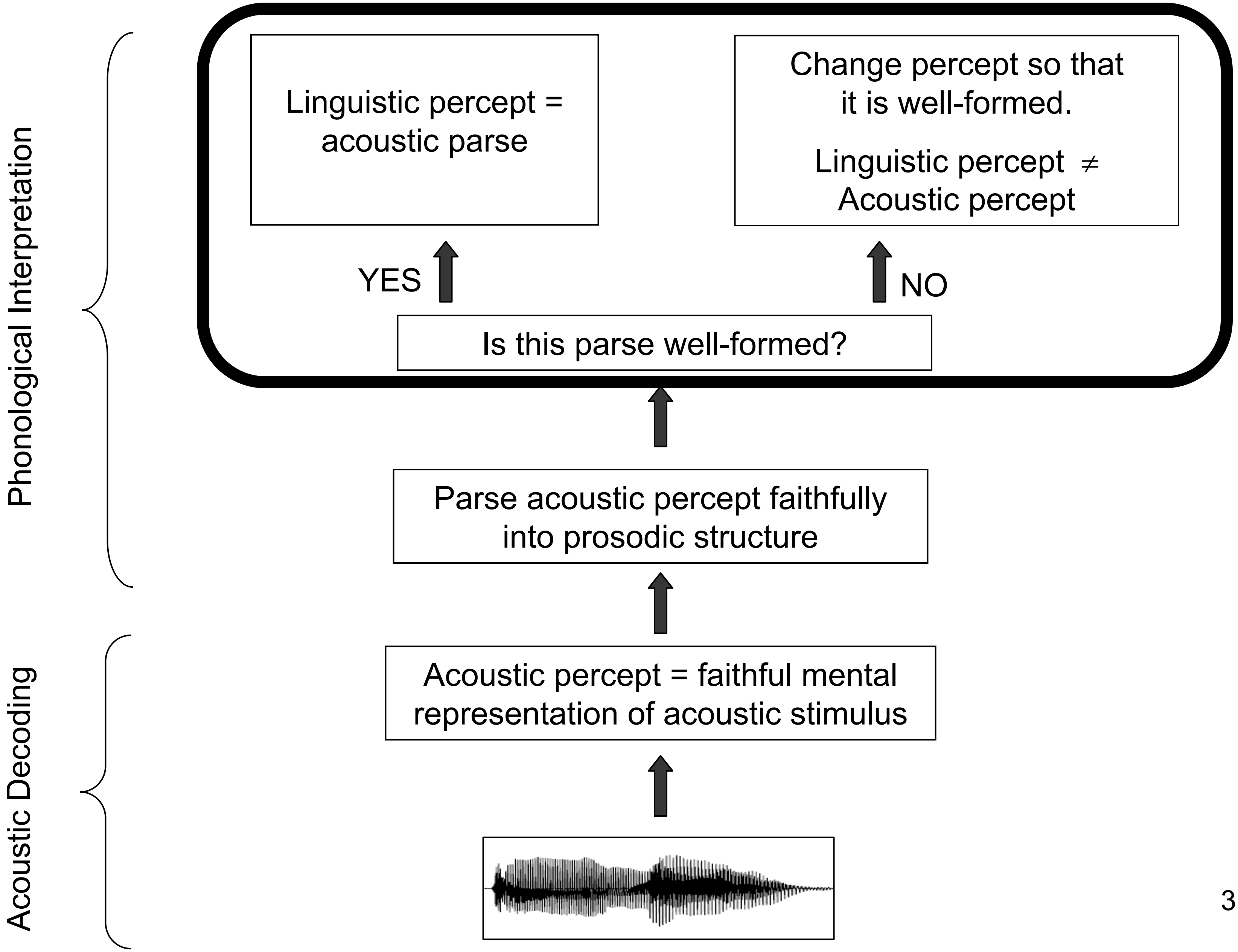
[k^h]ope vs. s[k]ope

[t^h]op vs. s[t]op

❖ We can use syllabically determined allophones like these to look for evidence that listeners parse the acoustic signal into syllables.

- Present listeners with an sound that can only appear in syllable initial/final position.
- Do listeners insert a syllable break before/after the sound?

Mapping from acoustic signal to linguistic percept



Perceptual epenthesis as evidence of syllabification

- When acoustic stimulus contains a sound sequence that is not well-formed, listeners perform perceptual epenthesis.
  - **Japanese** (Dupoux et al. 1999)
    - [ebzo]~[ebuzo] continuum.
    - French listeners perceived endpoints accurately.
    - Japanese listeners perceived [ebzo] endpoint as [ebuzo]

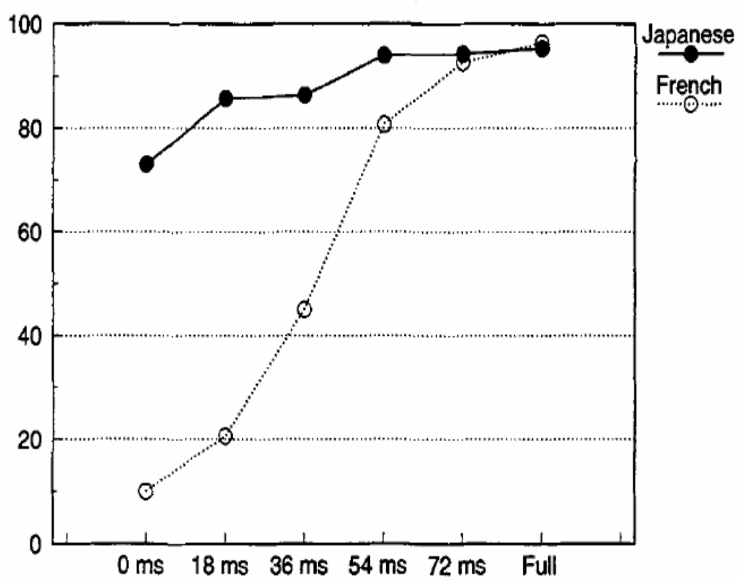


Figure 1. Percentage (*y* axis) of [u] vowel judgments in stimuli such as *ebuzo* in French and Japanese participants as a function of vowel duration (*x* axis) in Experiment 1.

- **Evidence for syllabification?**
  - Acoustic percept = [ebzo]
  - Prosodic parse = [eb.zo], [e.bzo]
  - Neither are well-formed.

    [b] only in onset in Japanese.

    No tautosyllabic clusters allowed.
  - Perceptual system "corrects" the acoustic mapping via epenthesis.
  - Linguistic percept: [e.b<u>u</u>.zo]
- See Berent et al. (2007) and Berent and Lennertz (2007) for similar evidence from English, and Kabak and Idsardi (2007) for Korean.

Perception of [s + stop]-sequences in English

$$/p, t, k/ \rightarrow \begin{cases} [p^h, t^h, k^h] \ / \ _{\sigma}[\ \_\_ \\ [p, t, k] \ / \ _{\sigma}[s\ \_\_ \end{cases}$$

- Aspirated stops occur *only* in syllable initial position.
  - If acoustic percept contains an aspirated stop, then only prosodic parses where this stop appears in syllable initial position will be well-formed.
  - If no such parse of the acoustic percept is possible, then the perceptual system needs to "correct" the acoustic percept with perceptual epenthesis.

			Perceptual hypotheses		
			Faithful to acoustic percept		**Perceptual epenthesis**
Acoustic percept			**Syllable break before stop**	**Syllable break before [s]**	**Syllable break before stop**
Aspirated	**Initial**	[spʰika]	s.pʰi.ka	spʰi.ka	sə.pʰi.ka
	Medial	[luspʰika]	lus.pʰi.ka	lu.spʰi.ka	lu.sə.pʰi.ka
Unaspirated	**Initial**	[spika]	s.pi.ka	spi.ka	sə.pi.ka
	Medial	[luspika]	lus.pi.ka	lu.spi.ka	lu.sə.pi.ka

Experiment 1: Same or Different?

- Recording: [lusəp^{h}ika].
- By splicing out [lu-], [ə] and [^{h}] in different combinations, created six stimuli.

		[+asp] stop	[-asp] stop
With schwa	**Medial**	lusəp^{h}ika	
	Initial	səp^{h}ika	
Without schwa	**Medial**	lusp^{h}ika	luspika
	Initial	sp^{h}ika	spika

Experiment 1: Same or Different?

- Created pairs of tokens
  - Identical = for four [+asp] tokens
  - Non-identical = one member always {[+asp], with schwa}
- Presented in random order to subjects in same/different task.
- Table below shows kinds of non-identical pairs, with expected response.

	Token 1	Token 2	Expected	Reason
[+asp]	lusəpʰika	luspʰika	Different	Legal parse for Token 2 possible.
[+asp]	səpʰika	spʰika	Same	No legal parse for Token 2. Perceptual epenthesis expected.
[-asp]	lusəpʰika	luspika	Different	Legal parse for Token 2 possible.
[-asp]	səpʰika	spika	Different	Legal parse for Token 2 possible.

Experiment 1: More on the design

❖ Token selection: 3 stops × 2 vowels

	Labial	Coronal	Dorsal
[i]	vɛsəp^{h}imu	kusət^{h}ila	basək^{h}ifa
	lasəp^{h}ika	fasət^{h}imo	masək^{h}ilu
[a]	fisək^{h}ana	lusət^{h}api	masəp^{h}ali
	visəp^{h}ano	misət^{h}aku	pisək^{h}ami

❖ Four blocks. Identical pairs once per block, non-identical twice per block

- Per block: 12 recordings × ((4 non-identical × 2) + (4 identical)) = 144.
- 4 blocks × 144 = 576 pairs.
- Takes roughly an hour to complete the experiment.

❖ Participants

- 15 Michigan undergraduates.

Experiment 1: Results

- Signal Detection Theory

		Actual	
		Different	**Same**
Response	**Different**	Hits	False alarms
	Same	Misses	Correct rejections

- p(H) = Hits / Different

  p(FA) = False alarms / Same

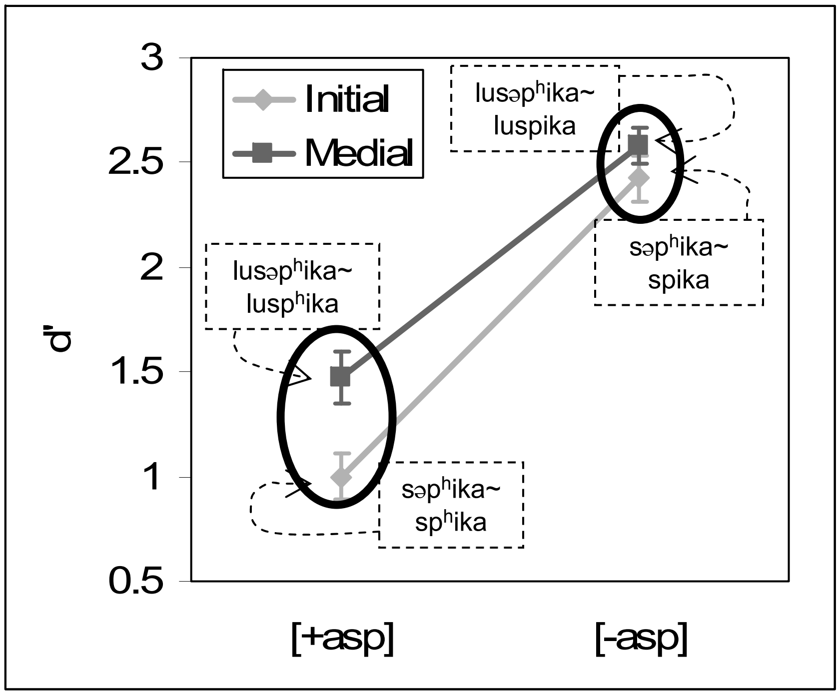
- d' = Norm(p(H)) – Norm(p(FA))
  - High d' ⇒ High discriminability
  - Low d' ⇒ Low discriminability (high confusability)

- Calculate d' for each token-type in each condition, for each subject.

Experiment 1: Results

d'-scores

Condition		Token 1	Token 2	d'	Expected
[+asp]	**Medial**	lusəpʰika	luspʰika	1.49	Different
	Initial	səpʰika	spʰika	0.99	Same
[-asp]	**Medial**	lusəpʰika	luspika	2.59	Different
	Initial	səpʰika	spika	2.44	Different



Repeated measures ANOVA

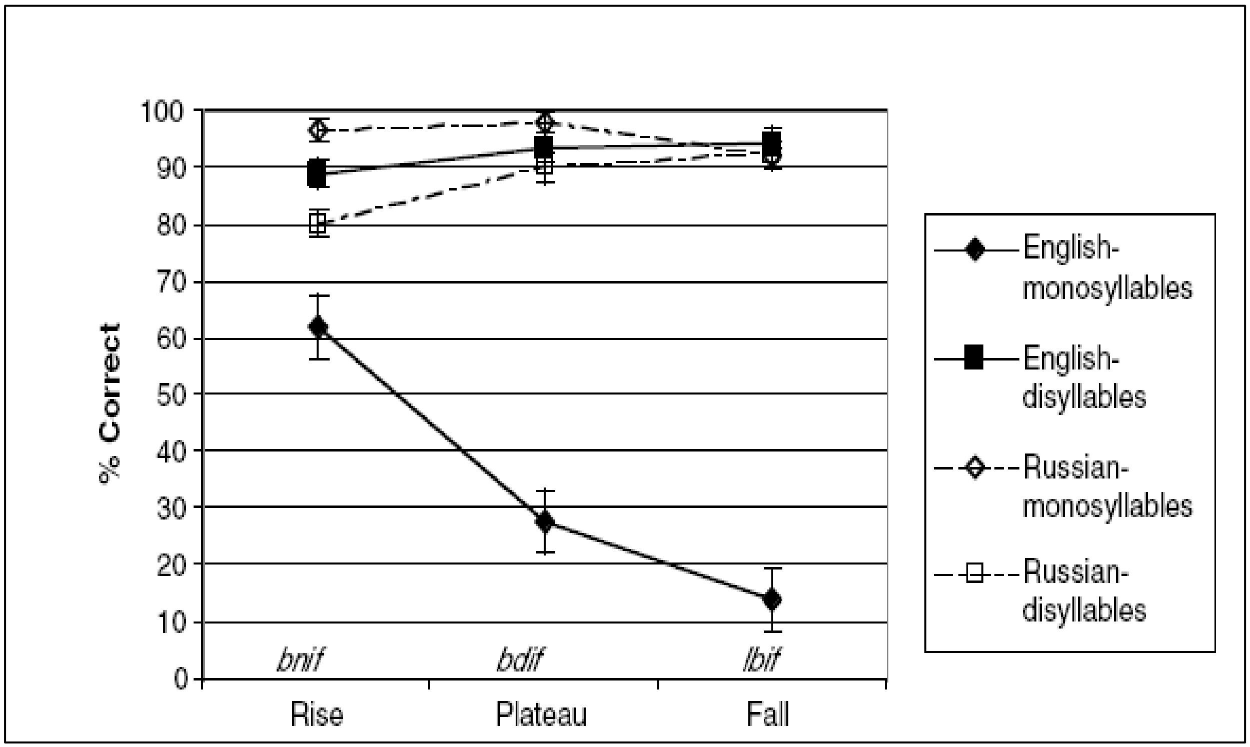
Factor	Subject		Item	
	F (1, 14)	*p*	*F* (1,11)	*p*
Aspiration	139.0	< .001	575.9	< .001
Position	15.7	< .002	73.4	< .001
Interaction	12.0	< .005	12.8	< .005

Experiment 2: Syllable count

❖ Inspired by Berent *et al*. (2007). They presented English and Russian listeners with clusters that are observed in Russian but not in English.

Task of the participants was to count the number of syllables.

English listeners consistently perceived an extra syllable.



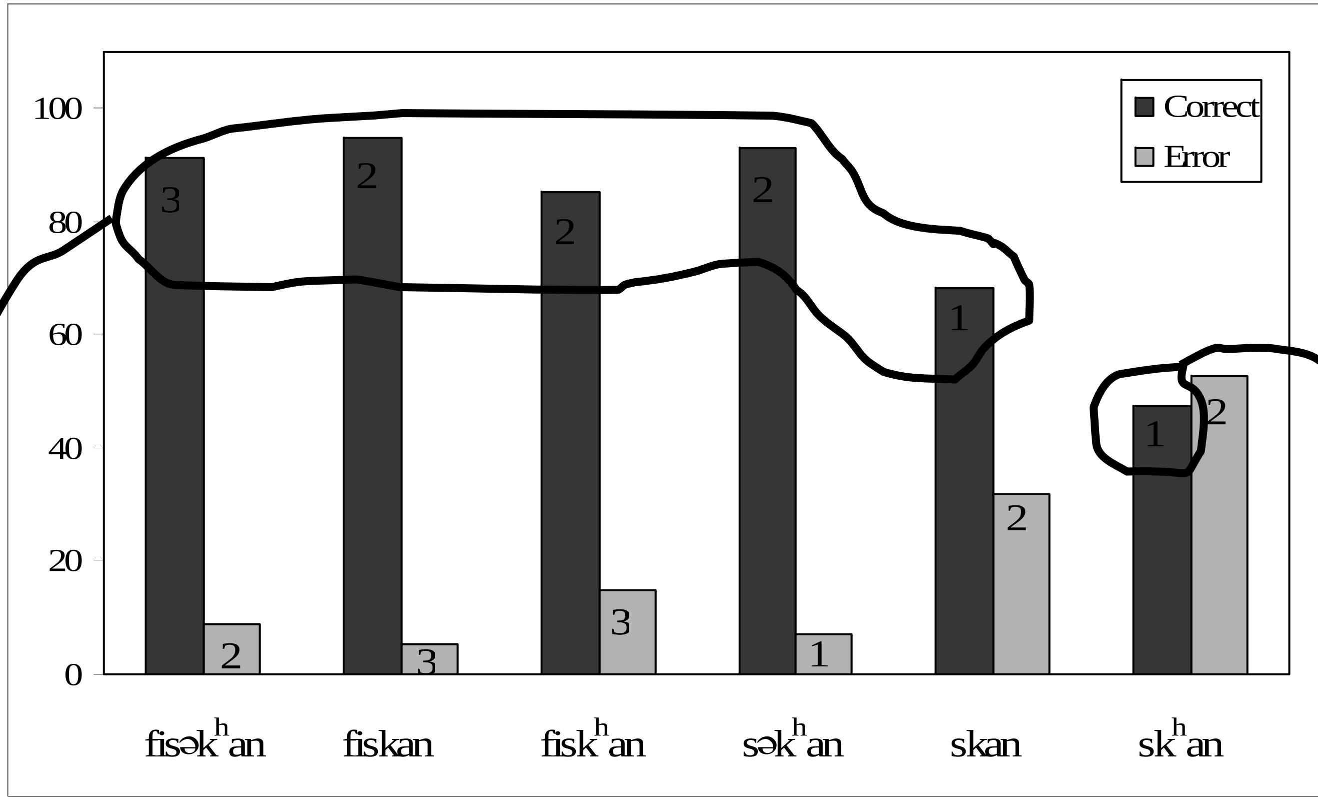
Experiment 2: Syllable count

- Six kinds of stimuli, created as in Experiment 1.

			Stimulus	Actual number of syllables	Expected response
Medial	With schwa	[+asp]	fisək^{h}an	3	3
	No schwa	[-asp]	fiskan	2	2
		[+asp]	fisk^{h}an	2	2
Initial	With schwa	[+asp]	sək^{h}an	2	2
	No schwa	[-asp]	skan	1	1
		[+asp]	sk^{h}an	1	2

- Four different token sets for each place of articulation. Each token presented 8 times.
- Also 60 fillers, each presented 4 times.
- Participants: 12 Michigan undergraduates.

Experiment 2: Results



Different from 50 at
p < .001 on one sample t-test.

Subject: $t(11) = .929$, $p = .78$
Item: $t(11) = .94$, $p = .37$

Discussion

❖ **Discriminability**

Medial		Initial
lusәp^{h}ika ~ lus$\underline{p}$ika	=	sәp^{h}ika ~ s$\underline{p}$ika
lusәp^{h}ika ~ lus$\underline{p^{h}}$ika	>	sәp^{h}ika ~ s$\underline{p^{h}}$ika

✓[spika] less likely to be identified as [sәp^{h}ika].

*[sp^{h}ika] more likely to be identified as [sәp^{h}ika].

❖ **Syllable count**

Above chance except for [sk^{h}an], where two syllable responses were about as likely as one syllable response.

❖ Perceptual epenthesis more likely in context where an acoustically faithful percept would be syllabically ill-formed.

- Somewhere in the mapping from acoustic percept to linguistic percept prosodic structure is imposed on the percept.
- This is sent through the grammar, and if necessary, the percept is altered to agree with the demands of grammar.

Discussion

- Hidden prosodic structure
  - [sp^{h}ika]/[sk^{h}an] contains no schwa after [s].

    But participants responded as if there was a schwa.
  - Mismatch between acoustics (physical properties of sound wave) and the response pattern of participants.
  - This can be interpreted as evidence that listeners impose hidden structures, such as syllables, on the acoustic stimulus somewhere during the mapping from acoustics to linguistic percept.
- Language users are conscious of allophonic distribution rules.
  - This determines not only production, but also perception.
  - They use this information to parse the acoustic signal into syllables.
  - Dupoux et al. (1999) and Berent et al. (2007) have shown that listeners use coarse grained phonotactics in this manner. But now we see that also more fine-grained distributional information can be used in the same manner.

Alternative explanation: transitional probability

- $p(p^h \mid \#s) = 0$; $p(p^h \mid ə) > 0$ } $p(\#səp^h) > p(\#sp^h)$

- One way to test this, in the spirit of Dupoux et al. (2001)
  - Assume that $p(p^h \mid i) > p(p^h \mid ə)$.
  - Discriminability should be even worse between [sip^{h}ika] and [sp^{h}ika].
  - More same responses for [sip^{h}ika]~[sp^{h}ika]
    Fewer same responses for [səp^{h}ika]~[sp^{h}ika]

- Either way:
  - Linguistic percept is adjusted to agree with linguistic expectation.
  - Whether this expectation is based on something like syllable structure.
  - Or transitional probability.
  - Transitional probability can also be incorporated under the umbrella of "grammar". See recent studies that gives grammatical explanations for "gradient phonotactics" (Anttila 2007; Coetzee & Pater 2006, to appear).

References

Anttila, Arto. 2007. *Gradient Phonotactics in Optimality Theory.* Ms. Stanford University.

Berent, Iris and Tracy Lennertz. 2007. What we know about what we have never heard before: Beyond phonetics: Reply to Peperkamp. *Cognition*, 104:638-643.

Berent, Iris, Donca Steriade, Tracy Lennertz and Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104:591-630.

Coetzee, Andries W. and Joe Pater. 2006. *Lexically Ranked OCP-Place Constraints in Muna*. Ms. University of Michigan and University of Massachusetts. [ROA-842.]

Coetzee, Andries W. and Joe Pater. to appear. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory*.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25:1568–1578.

Dupoux, Emmanuel, Christophe Pallier, Kazuhiko Kakehi, and Jacques Mehler. 2001. New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, 5(16):491-505.

Kabak, Boris and William Idsardi. 2007. Speech perception is not isomorphic to phonology: The case of perceptual epenthesis. *Language and Speech*.